ABSTRACT

A method for predicting an amino acid sequence compatible with a three-dimensional (3D) structure of a protein. A reduced virtual representation of the 3D structure is constructed, and, for each position along the representation, its solvent accessibility is determined. For each position along the structure, an amino acid residue is randomly selected from a predefined group of amino acids having a solvent accessibility compatible with the solvent accessibility of the position. A Monte-Carlo simulation is performed on this devised protein in which an amino acid at a particular position is sequentially replaced with other amino acids having the same solvent accessibility, and an energy score is calculated for each rotamer. The lowest scoring rotamer for this position is then selected The Monte-Carlo simulation is repeated for each position along the sequence, to obtain an amino acid sequence with the lowest total energy score.